DOCUMENT RESUME

ED 237 932 CS 007 375

AUTHOR Meyer, Linda A.
TITLE Long-Term Academic Effects of Direct Instruction Follow Through. Technical Report No. 299.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Nov 83
CONTRACT 400-81-0030
NOTE 47p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Basic Skills; Comparative Analysis; *Disadvantaged Youth; *Educational Research; Inner City; Longitudinal Studies; Primary Education; *Program Effectiveness; Reading Instruction; *Reading Programs; Reading Research; Remedial Programs; Teaching Methods
IDENTIFIERS *Direct Instruction; *Project Follow Through

ABSTRACT

The Direct Instruction Follow Through program of enhanced instruction for children in kindergarten through grade 3 was introduced in 1968 in schools serving disadvantaged students in 180 communities throughout the United States. To determine the long-term effects of the program, a study compared the performance of the first three cohorts of Follow Through children from P. S. 137, the Bainbridge School, in Brooklyn, New York, with that of a control group of non-Follow Through students from a school in the same area. Data were collected from school records for 82% of the Direct Instruction Follow Through students and 76% of the control group students. Specific information gathered included high school graduation date, ninth grade reading and math scores, and student's application to and acceptance by colleges. Data analysis revealed that (1) over half the Follow Through students finished high school as compared to just over a third of the control students, (2) Follow Through students dropped out significantly less than did control students in two of the three cohort groups, and (3) more Follow Through students applied for and were accepted by colleges. A comparison of the subjects' ninth grade performance indicated that Follow Through students did better than their control group counterparts in both reading and math. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED237932

CENTER FOR THE STUDY OF READING

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Technical Report No. 299

LONG-TERM ACADEMIC EFFECTS OF DIRECT INSTRUCTION FOLLOW THROUGH

Linda ... Meyer

University of Illinois at Urbana-Champaign

November 1983

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02238

The research reported herein was supported in part by the National Institute of Education under Contract No. NIE-400-81-0030. Ronald P. Toby helped conceive this study. Then, in order of appearance, Joan Gutkin, Zenith Terry, Carol Amazon, and Russell Gersten helped to trace children from their elementary schools to their intermediate schools, and finally to their 65 respective high schools. Thomas Minter, former Deputy Chancellor for Instruction of the New York Public Schools, his colleague JoAnn Asciuto, and most recently Anthony J. Alvarado, Chancellor of the New York City Public Schools worked with me to retrieve data. Russell Gersten and Bill White completed a preliminary data analysis. Comments on the manuscript from Ronald P. Toby, J. McVicker Hunt, John S. Ory, Thomas L. Good and the three reviewers strengthened this manuscript. The author fondly acknowledges the encouragement and support of the late Rachal Jean Mitchell, P.S. 137 parent, and Community School President, whose unflagging commitment to "our children" and their education helped to sustain Follow Through for more than a decade. This article is a token offering in her memory. Forthcoming, The Elementary School Journal.

CS 007375

EDITORIAL BOARD

William Nagy
Editor

R. Logan Bickford
Harry Blanchard
Nancy Bryant
Pat Chrosniak
Avon Crismore
David Dunning
Linda Fielding
Dan Foertsch
Meg Gallagher
Paul Hardin
Patricia Herman
Gabriella Herman
Asghar Iran-Nejad
Steve Levine
Margie Leys
Paul Mayberry
Mary Moes
Carol Peterman
Michael Reddix
Theresa Rogers
Judith Scott
Ileana Seda-Santana
Fernando Senior
Marcy Stein
Janice Stewart
Behrooz Tavakoli
Paul Wilson

Long-term Academic Effects of Direct Instruction Follow Through

It has been fifteen years since the Follow Through program of enhanced instruction in early childhood education was introduced for grades kindergarten through three in elementary schools serving disadvantaged children in 180 communities throughout the United States.[1] Children who were in the first cohort of Follow Through classes are now about twenty years old--old enough to be sophomores in college, high school and junior high school dropouts, shopclerks or shop mechanics. This means that we now have our first opportunity to measure the long-term effects of Follow Through instruction through several of the most important career-branching points in a young person's life. Did they finish school or drop out? Did they apply to and enter college? How was their academic performance in high school? In short, we now have an opportunity to test Follow Through as a long-term social and educational investment, to see whether it has paid off, and to suggest implications for early childhood and other educational programs. The primary research question was how the Follow Through graduates compared through high school with a control group. Of secondary interest was how the Follow Through students' third grade performance correlated with their ninth grade performance in reading and math.

In this paper I shall compare the long-term performance of the first three cohorts of Follow Through children from P.S. 137, the Bainbridge School, in Brooklyn's Ocean Hill - Brownsville

section, with a control group of non-Follow Through children from a school in the same area just a few blocks away. There are just a few longitudinal studies of disadvantaged children who have been in early childhood programs. The results of these investigations are reported next, though none of the studies reports on the high school performance of the early childhood program graduates.

Hunt (1982) synthesized almost a hundred studies on developmental theory and early childhood interventions that resulted in higher performance for predictably low-performing children. He declared, "The fiction that psychological development has a fixed rate predetermined by heredity is hard to reconcile with the findings," (Hunt, 1982, p. 5). Hunt's research focuses on intervening with infants at birth, and their mothers, or foundlings such as children in the Tehran Orphanage and their caregivers. In all of these settings, Hunt's intervention strategies produced demonstrable differences favoring the experimental children-- those children receiving instruction. These studies strongly suggest that early intervention will benefit infants and young children in high risk situations--those children most likely later to fail in school.

Lazar, Hubbel, Murray, Rosche, and Royce (1977) and then Darlington (1981) studied fourteen infant and preschool experimental programs to search for common effects from early childhood programs. They report generally positive treatment

effects with fewer experimental students placed into Special Education classes. Only 8.6% of the experimental students, in contrast to an average of 29% of the control students, required special education placement. This report, most often referred to as "The Consortium," also found significant differences in retention. Twenty-four percent of the control students were retained, but fewer than 19% of the experimental students were held back. It is interesting to note, though, that among the fourteen experimental groups with the highest retention rate (almost 50%) also showed the lowest percentage of students later placed into special education classes. Thus, these findings suggest that programs that retain students as a preventive measure--using the holdover time as an opportunity for the students to catch up before they are promoted--may have long-term benefits. These findings may raise doubts about the validity of using retention as a measure of long-term effectiveness in early childhood programs, since it appears that early retention may increase the chances that low-performing students will do well later in school.

The Consortium studies also report IQ scores for the experimental and control students. The later effects of these early childhood programs when IQ was measured immediately, and then again after one, two, three, and four years, show an average drop of about 4 more IQ points for the control students than the experimental students. The remaining data in these reports are

from student or parental interviews dealing with topics such as how individual students think they are doing in school, or whether the parents feel the preschool program was good for their child. Therefore, these data are difficult to quantify and interpret.

Another report focuses on just one of the fourteen programs reported by the Consortium. It is, Young Children Grow Up: The Effects of the Perry Preschool Program on Youths Through Age 15, by Schweinhart and Weikart (1980). This book describes Weikart's Cognitive Curriculum which draws heavily from the teachings of Piaget, and the long-term effects of this program on young children. The major findings from Weikart's laboratory work in the Perry Preschool Program are:

1. There was a difference of about 12 IQ points for the experimental and control groups at age 4, but the differences in IQ disppeared and both groups had IQ's of about 81 at age 14.

2. Weikart's group showed more motivation and potential in elementary school.

3. The experimental group placed higher values on schooling at age 15.

4. Norm-referenced achievement test data, reported as percentages of items passed by each group on the California Achievement Test (the CAT), show the experimental students outperformed the control group (54% to 47%) at age 8. At age 14, the experimental group dropped to 36% correct and the control

group to 28% items passed for Total Achievement. Reading performance has the experimental group at 62% and the control group at 56% at age 8, but at 37% and 30% respectively at age 14. Math and language performance are similar, although the experimental group at age 14 outperforms the control group 39% to 29% items correct.

5. The Perry Preschool study also reports that "deviant behavior" in school was less for the experimental group. The experimental group was more self-confident, although the differences were not great.

The Consortium study and the Schweinhart and Weikart work ended before the students finished high school. But, the drop in IQ for the Schweinhart and Weikart experimental subjects, and similar drops by the students in the Consortium group, show that major differences between these two groups have diminished by the beginning of high school. With both of these studies, the longer one traces effects, the fewer differences there are between the experimental and control groups.

This study extends the findings from the previously cited studies of the long-term effects of early childhood education programs and departs from those studies in five ways. First, it reports on students who had been in a very structured kindergarten through third grade early childhood program, Direct Instruction Follow Through, as developed by Engelmann and Becker in 1968. The other longitudinal studies are of much more

traditional, therefore, less academic and shorter term early childhood programs. Second, this is a study of the long-term effects of a school-based intervention. Therefore, the results of this work have ecological validity and significance for early childhood education in school settings. Third, all data were collected in naturalistic settings, from school records in regular schools in New York City, whereas the Schweinhart and Weikart students were in a laboratory program as were the Consortium Study students. Fourth, in this study, I followed Direct Instruction Follow Through graduates and a control group through high school, thereby completing the measures of elementary and secondary school success. And, last but perhaps most important, this study focuses on general outcome measures of schooling--graduation and college application, as well as upper grade reading and math achievement. I agree with Maeroff (1982) that "A diploma is no longer a luxury; it is the key that opens the first lock on the door to the job market" (p. 47).

In this study, I trace the performance of Follow Through graduates (children who had either three or four years in the Direct Instruction Program at P.S. 137) to see if there appear to be lasting effects in high school from their early childhood Follow Through program. P.S. 137 Follow Through graduates were selected for this study because this school: (a) was the only one of the New York City and Philadelphia inner-city Follow Through sites to demonstrate significant student achievement gains and

positive affective outcomes (Stebbins, St. Pierre, Proper, Anderson, & Cerva, 1977; p. 150) at the end of third grade; (b) has continued to demonstrate its effectiveness in basic skills for over a decade (Meyer, Gersten, & Gutkin, 1983); and, (c) is characteristic of inner-city schools serving disadvantaged minority students. Therefore, a program of demonstrated long-term effectiveness in this setting has implications for other inner-city elementary school programs.

As I shall detail later, P.S. 137 and the control school were fairly comparable in racial composition and very comparable in economic composition and reading achievement prior to the implementation of Follow Through in 1968. Both schools draw from very similar neighborhoods of multiple family row houses, abandoned buildings, vacant lots, and corner stores. This area shows many signs of racial isolation and poverty.

METHOD

Setting

The two schools in this study are in District 23, the Ocean Hill-Brownsville section of Brooklyn. This area is well known for being one of the most disadvantaged areas of the country (New York Times, February 4, 1980). District 23 is one of the lowest-performing of the 32 school districts in New York City. The fifteen elementary schools in the district average 519th in rank out of the 630 elementary schools in New York City.

P.S. 137 is the only Direct Instruction Follow Through school in New York City. Thus, it was the obvious choice for this study. We chose a second school in District 23 that we believed to be demographically equivalent to P.S. 137 to serve as the control school. New York City Central Board of Education records going back to 1968 revealed that the two schools had similar patterns of reading achievement (in 1966-1967), ethnicity (in 1969-1971), and socioeconomic status (in 1973). Table 1 presents the comparisons for the two schools. In both schools, students in second and third grade entered these grades at comparable levels (1.6 at second grade and 2.4 at third grade in grade equivalents) on the Metropolitan Achievement Test for the years immediately preceding the implementation of Follow Through. Likewise, these students completed second and third grade at comparable levels (2.3 and 2.5 at the end of second grade; 3.2 and 3.3, respectively at the end of third grade) during those years.

Insert Table 1 about here.

The two schools then had roughly equivalent Black and Puerto Rican populations in 1969, 1970, and 1971-the first three years of Follow Through. P.S. 137 averaged over 85% Black students (84%, 87.3%, and 89.3%), whereas the control school's Black population was approximately 82% (82.2%, 80.7%, 82.3%) for the

same three years. The Puerto Rican populations shifted during the 1969-1971 school years with a little over 11% (14.3%, 10.3%, 9.2%) of the P.S. 137 children from Puerto Rican heritage, compared to about 17% (17.1%, 18.4%, 16.3%) from the control school.

The 1973 percentages (in 1973 Cohort 2 Follow Through students were in third grade) from the two schools for families receiving aid for dependent children (AFDC) were also very similar. Almost 77% (76.9%) of the P.S. 137 families received AFDC in 1973, compared 77.5% of the families from the control school. Thus, the AFDC percentages were still nearly equivalent when the Cohort 2 Follow Through students were in third grade. Therefore, we demonstrated that the two schools had comparable reading levels before Follow Through began and comparable ethnic mixes and percentages of families receiving support for dependent children during the implementation of Follow Through.

Treatments

The Direct Instruction Follow Through Model at P.S. 137 attributed its results to: (1) sponsorship--the school worked closely with Project Manager appointed by the University of Oregon to transmit the Direct Instruction Model to the site; (2) the Distar (Engelmann & Stearns, 1972) curricula in reading, language, and math with its emphasis on carefully sequenced lessons, high amounts of practice, and mastery; (3) increased allocated time to basic skills instruction--about 60 percent of

the school day is spent on reading, language, and math; (4) monitoring of criterion-referenced and norm-referenced student progress; and (5) teacher training in workshop and classroom settings to focus on student performance (Meyer, et al., 1983).

At P.S. 137, the Follow Through program begins in kindergarten, but students may also enter as first graders. Regardless of whether children begin the program in kindergarten or first grade, they continue through three levels of the program (see Becker, 1977; or Meyer, et al., 1983, for more thorough descriptions of the Direct Instruction model). Because of these two entry points and the model's emphasis on acceleration based upon skill mastery, the highest-performing groups of children may finish all three levels of the reading, language, and math programs by the end of second grade. Low-performers, on the other hand, may not finish the third level programs by the end of third grade. Once groups finished the Direct Instruction programs, in second or third grade, they went into the basal reading and math series adopted by the school.

The students in the control school were in traditional classes that emphasized reading and math instruction starting in first grade.

Finding the Students: Procedures

The Direct Instruction Follow Through Students. The first step was to identify all children in the first, second, and third

cohorts (groups entering Follow Through in the 1968-1969, 1969-1970, or 1970-1971) school years. In the spring of 1981, when the Cohort 1 students were seniors in high school, the Follow through Coordinator and I began to search the Follow Through records to make up rosters of the Cohort 1, 2, and 3 students. We had to reconstruct the rosters since there had been no on-going contact with these students since they graduated from Follow Through at the end of third grade. We then reduced the rosters to eliminate students who had not remained in the Follow Through program for either three or four years. This study is limited to three and four year students because we felt the truest test of the model's effectiveness would be with students who had completed three levels of the program. We sought to follow three cohorts of students so that each cohort could replicate the findings from the other cohorts, as suggested by Cook and Campbell (1979).

The second step was to trace the Cohort 1, 2, and 3 students to their respective high schools. The procedures for tracing each cohort are slightly different, so we will report the procedures sequentially. We began tracing Cohort 1 and 2 students in the spring of 1981, working fror the Follow Through records and the school records in the control school to find the transfer school for each child on our rosters. We searched the records by hand, recording the intermediate school the student reported that he/she planned to attend.

Next, we organized and coded the lists of students by school and contacted the transfer school either by going to the school if there were a number of students identified for that school, or by telephone if we had fewer students listed. If that intermediate school had received, but later discharged the student, we contacted the "new" intermediate school. We continued this process until either we had the name of each student's high school or until we could no longer trace the student. In addition, we wrote letters to parents, interviewed former classmates, and questioned paraprofessional classroom aides who had had children in the first two cohorts about students that we could not locate, thereby utilizing several means to locate as many students as possible. There was one wave of data collection for Cohort 1 because there are no computer records on those students. We had two waves of data collection for the Cohort 2 and 3 students because staff from the New York City Central Board of Education, in the Office of the Deputy Chancellor for Instruction, did searches of their computer files. Additional Cohort 2 and 3 students were found with the computer search, so we did a mailing to high schools for those students' records. All data reported come from school records.

Instrument. Each student's name and date of birth appeared on an individual form. Students were not identified on these forms as members either of the Direct Instruction Follow Through

group or the control group. The form asked for the following information:

(1) high school graduation date

(2) ninth grade reading score

(3) ninth grade math score

(4) student's application to college

(5) student's acceptance to college

(6) student's special education placement

(7) student's school attendance for the previous year

Approximately two weeks after the forms were mailed, we began to telephone and visit schools that failed to respond to the letter and questionnaire. We continued this procedure for approximately three months. Then, to pursue as much of the missing information as possible, we again telephoned or visited schools to gather the remaining data.

During the visits and follow-up phone calls, we pressed for all of the information on each student. We took particular care to define the term, "transferred." Occasionally, students' records were coded "17+." This code sometimes accompanied "transferred," on the student record folder. We found that "17+" meant either the student was over 17 and had dropped out of high school, or it meant the student had transferred to another school. School personnel are understandably sensitive about the high percentage of inner-city students who drop out and fail to finish high school (45-50% New York Times, March 15, 1983 or up

to 68%, New York Times, June 10, 1983). Therefore, we took care to explain that we were looking for all long-term effects from an early childhood program, and that while it was important to keep track of dropouts, we were interested in various information on each student. Only if the second school contacted yielded no information, were students considered lost and subsequently dropped from the study.

Results

Percentages graduated, retained, dropped out, applied to college, and accepted for college. Table 2 summarizes the results of this portion of the study.

Insert Table 2 about here.

Sixty-two percent of the Direct Instruction Cohort 1 students, sixty-four percent of the Cohort 2 students, and sixty-two percent of the Cohort 3 students graduated from high school, whereas only 38% of the Cohort 1 and 2 Control students and 39% of the Cohort 3 Control students graduated. A little over 4 percent of the Direct Instruction Cohort 1 students, fifty percent of the Cohort 2 students, and 10 percent of the Cohort 3 students were held back at least one grade. Approximately forty-two percent of the control group Cohort 1 students, thirty-five percent of the Cohort 2 students, and twenty-one percent of the

Cohort 3 control students repeated at least one grade. The dropout rates vary by cohort, though for the first two cohorts, substantially more control students than direct instruction students failed to complete high school. Thirty-nine, twenty, and twenty-four percent, respectively, of the Cohort 1, 2, and 3 Direct Instruction students dropped out of school, whereas, sixty-two, fifty-eight, and eighteen percent of the control group students dropped out.

Half of the Direct Instruction Cohort 1 students applied to college, and 41 percent were accepted; twennty-three percent of the Cohort 2 students applied and were accepted; and thirty-eight percent of the Cohort 3 students applied and were accepted for college. In contrast, eighteen percent of the Cohort 1 control group, twenty-five percent of the Cohort 2 control group, and twenty-four percent of the Cohort 3 control group applied to college. Nine percent of the Cohort 1 control group, twenty-one percent of the Cohort 2 control group, and twenty-one percent of the Cohort 3 control group were accepted for college.

The results of the t-tests for the percentages graduated, retained, dropped out, applied to and accepted for college, aggregated for Cohorts 1, 2, and 3, Direct Instruction and Control students appear in Table 3. All differences are

---

Insert Table 3 about here.

---

significant beyond at the $p = .01$ level. Approximately sixty percent of the Direct Instruction students graduated from high school in comparison to almost thirty-eight percent of the control students. About twenty-one percent of the Direct Instruction students and almost thirty-three percent of the control students were retained. A little under twenty-eight percent of the Direct Instruction students have dropped out of high school, whereas almost half (46%) of the control students have left school. Almost twice the percentage of Direct Instruction students have applied to college (34%) in comparison to eighteen and a half percent of the controls. Exactly twice the percentage of Direct Instruction students were accepted for college (34%) as control students (17%).

Ninth grade reading and math scores. The ninth grade reading and math score comparisons appear in Table 4. All scores are reported in grade equivalents because these are the only scores available. The ninth grade scores are Total Reading or Total Math scores on the California Achievement Test. The Direct instruction students scored 9.6 (SD = 2.22), 8.51 (SD = 2.16), and 9.[illegible] (SD = 2.24) in reading for the three cohorts. The control groups scored 7.59 (SD = 2.54), 7.96 (SD = 1.92), and 8.94 (SD = 2.59) in reading for the same three cohorts. The aggregated reading score for the Direct Instruction students is 9.20 (SD = 2.22), and 8.21 (SD = 2.42) for the control group. The t-tests for the Cohort 1 comparisons and for the aggregated

comparisons are significant at the $p = .01$ level. Cohort 1 Direct Instruction students averaged 8.3 (SD = 2.13) in math, Cohort 2, 8.08 (SD = 2.52), and cohort 3, 9.03 (SD 2.35). The control students averaged 7.79 (SD = 2.67) for Cohort 1, 7.60 (SD = 1.79) for Cohort 2, and 8.33 SD = 2.23) for Cohort 3 in math. The aggregate of the three cohorts shows the Direct Instruction groups averaging 8.59 (SD = 2.33), and the control students 7.95 (SD = 2.23) in math.

Insert Table 4 about here.

Correlations of third grade reading and math scores with ninth grade scores. End of third grade reading and math scores were available only for the Direct Instruction students. We calculated Pearson Product Moment correlations for third grade Wide Range Achievement Test (the WRAT) Reading and Math; Slosson IQ; and Metropolitan Achievement Test (the MAT) reading scores with ninth grade reading and math scores for Follow Through graduates. These correlations appear in Table 5.

Insert Table 5 about here.

There are generally high correlations for MAT and WRAT end-of-third grade reading scores and ninth grade reading. Third grade MAT total reading scores and ninth grade total reading

scores correlate .78 for Cohort 2 and .81 for Cohort 3. The end-of-third grade WRAT Reading scores and 9th grade reading scores correlate .47 for Cohort 1, .71 for Cohort 2, and .73 for Cohort 3. End-of-third grade WRAT math scores correlated with 9th grade math scores .49 for Cohort 1, and .39 for Cohorts 2 and 3. Slosson IQ scores were available only for Cohorts 1 and 2. Correlations of 9th grade reading scores with Slosson IQ scores were .37 for Cohort 1 and .49 for Cohort 2. Ninth grade math scores and Slosson IQ scores correlated .20 for Cohort 1, and .49 for Cohort 2.

To determine the equivalence of the sample of students found in this longitudinal study to the entire group of Direct Instruction students completing third grade, we calculated the sample's end-of-third grade reading performance and compared it to the data available on the entire cohort's reading performance and to the performance of poor students only in each cohort. The average performance of each cohort found was lower at the end of third grade in reading and math on the WRAT and the MAT than the average performance of either the total cohort or the poor children only subsample of the cohort.

We seldom received information requested for Special Education placement and attendance. Therefore, those data are not reported.

Discussion

In this part of the paper, I will discuss the results of this longitudinal study of Direct Instruction Follow Through students in terms of: (1) the Cohort 1 and 2 and 3 students that we traced successfully through high school; (2) the percentages of graduates, retainees, dropouts, college applicants; and college acceptances; (3) a comparison of ninth grade reading and math scores for Follow Through graduates and the control students; and (4) correlations of the end-of-third grade and ninth grade norm-referenced scores for Follow Through graduates. The final segment of this paper will present the implications of this study and suggestions for future research.

We found 82% of the Direct Instruction Cohort 1, 2, and 3 students. In comparison, we found 76% of the control students. While some might ask if we found students who are representative of both groups, we can argue that if anything, our sample of Direct Instruction students performed slightly lower at the end of third grade than the entire group of Direct Instruction Follow Through students performed at that time. Also, the identification process was the same for P.S. 137 and the control school when we developed the cohort rosters. Our data retrieval process was such that no one involved in collecting data could identify students by early childhood program, thereby biasing the results by pursuing some students more vigorously than others. The 16% attrition rate is actually quite low for a district whose

school rosters turn over on an average of 40% each year. It is safe to assume, therefore, that we found similar students from both P.S. 137 and the control school, though these may be the most stable students from both schools.

Percentage of Graduates, Retainees, Dropouts, College Applicants, College Acceptances, and Graduates. In all cohorts, over half of the Direct Instruction Follow Through graduates finished high school as compared to just over a third of the control group students. These are statistically significant differences that point rather clearly to apparent later effects from the Direct Instruction Follow Through program. Children from Direct Instruction Follow Through graduate from high school at a higher rate than students from the control school, although a slightly higher percentage of Direct Instruction Cohort 3 students dropped out of high school.

Some of the differences in graduation rates may be attributed to the higher percentages of Puerto Rican students in the control school (P.S. 137 averaged 11% Puerto Rican students while the control school averaged 17% Puerto Rican students). Dropout rates are highest in New York City for Hispanic students with 36% of those 18-19 year old students dropping out (New York Times, June 20, 1982). Forty-six percent that dropped out from the Cohort 1, 2, and 3 control groups is still much lower than the national figure of 73% Hispanic dropouts in the United States as a whole (Grant & Eiden, 1982). Thus, the District 23 students

from both elementary schools are graduating at higher than expected national averages for their ethnic groups.

Retainees. A little over 4% (4.3%) of the Direct Instruction Cohort 1 students and 42% of the control group were held back. Cohort 2 percentages are appreciably higher for both groups. We can document that the principal at P.S. 137 held over all students not at grade level in reading in the early 1970's. This would explain at least some of the jump from 4.3% to 50% of the Direct Instruction students retained. We have no comparable explanation for the drop from 42% to 35% of the control group students retained from these two Cohorts. Both groups have fewer students dropping out from Cohort 3, 10% Direct Instruction, and 21% controls, though these percentages may increase if some students who are still in school from Cohort 3 fail to graduate this year.

The inconsistency of the retention percentages and subsequently, the very low percentages of students from those Cohorts dropping out of high school support the Consortium (Lazar, et al., 1977) findings for one of its fourteen treatments where the program holding back almost 50% of its students later had the lowest percentage of students placed into Special Education classes. The findings from this study also suggest that retaining lower-performing students in the early grades may help them to succeed in the later grades.

Dropouts. It is impossible to look at the percentages of students graduating from high school without looking simultaneously, of course, at the percentages of students dropping out of high school before graduation, though some students in this study are still in school. The Direct Instruction Follow Through graduates dropped out significantly less than the control students from Cohorts 1 and 2. While some could argue that the difference in the percentage of Hispanic students at P.S. 137 and the control school clouds the dropout effects for these groups, there are such clear differences favoring the Direct Instruction Cohort 1 and 2 students that it is very likely that program effects account for significant amounts of the difference. Despite the higher percentage of Direct Instruction Cohort 3 students dropping out (24% to 18%), when aggregating, the 3 cohorts, 22.9% variance for dropouts is accounted for by treatment.

College applicants. The percentages applying to and being accepted for college present another comparison of the two groups. While there are clear effects for the Direct Instruction Cohort 1 and Cohort 3 students, there are no such effects for the Cohort 2 groups. When aggregating the three cohorts for each group, 53% of the variance for college application is accounted for by treatment. Likewise, 58.8% of the variance for college acceptance is accounted for by treatment.

If we look across the five general measures of high school success in this study, the pattern of effects for all three cohorts clearly support the Direct Instruction Follow Through graduates. The Direct Instruction students perform significantly better on twelve of the fifteen measures when we test each cohort separately. When we aggregate the three cohorts, Direct Instruction students perform significantly better than their controls on all five measures. Higher percentages of those students graduated from high school, fewer dropped out of high school, more applied and were accepted for college. Equally interesting, but less strong are the comparisons of ninth grade reading and math scores and the correlations of third grade and ninth grade reading and math scores reported next.

Ninth grade reading comparisons. There are significant differences in ninth grade reading for Cohort 1 students and less strong, though consistently positive effects favoring the Direct Instruction students in all 3 cohorts for reading and math. It is possible that the implementation of Distar Reading III (Engelmann & Stearns, 1972) which was in a field tryout edition for the Cohort 1 and 2 students accounts for some of the high reading performance for these students. The Distar Reading III materials rely extensively upon expository passages written for the program from which the students learn a number of scientific rules about the three forms of matter, momentum, and skeletal system, to name just a few. There is now substantial research

(Durkin 1978-1979; Durkin, 1980; Mason, 1983) to document that there is very little narrative or expository reading comprehension instruction in traditional classrooms. It is possible, therefore, that the combined effects of the lack of reading comprehension instruction for the control groups and the heavy emphasis on reading expository text in Distar Reading III for the Direct instruction students contributed these differences. It is also likely that the Distar Language instruction that paralled the students' reading instruction but focused on oral problem solving and other skills contributed soundly to this long-term reading performance.

Ninth grade math comparisons. The math differences also favor the Direct Instruction groups for all cohorts. There is a 5-month difference in performance for Cohort 1 students, about a 4-month difference for Cohort 2 students, and an almost 7-month difference for Cohort 3 students. It is popssible that math performance in ninth grade is more dependent upon math instruction beyond third grade than is reading instruction. Although Direct Instruction students outperform their control groups in math, they are still about half a year below grade level in ninth grade. This performance suggests that both groups would benefit from additional math instruction. The ninth grade comparisons in reading and math differ somewhat from the results for Direct Instruction graduates in fifth and sixth grade reported earlier by Becker and Gersten (1982). The Becker and

Gersten work showed consistent strong effects for WRAT reading and math problem solving and spelling, but only moderate effects in Total Reading and Total Math on the Metropolitan Achievement Test. The Becker and Gersten results for MAT Total Reading and Total Math and Science show lesser effects than the results of this study. Apparently, there are stronger long-term effects in Total Reading in ninth grade than there are in fifth or sixth grade. There are three possible explanations for the difference in these results. First, continuous effects in Total Reading may favor four-year Direct Instruction implementations--those sites where the program begins in kindergarten and continues through third grade. The Becker and Gersten (1982) study is limited to first grade starting sites. The extra year of treatment could account for the difference. Second, P.S. 137 and this control group are more disadvantaged than the sites studied by Becker and Gersten (1982). Perhaps the Direct Instruction program's most potent effects are in the most disadvantaged communities--those communities where schools have the greatest impact. The third possible explanation is that the Direct Instruction effects simply are less pronounced in the middle grades but then rebound in early high school and strengthen through high school to cause greater end of high school differences. One could test all three of these hypotheses by replicating this longitudinal study in a variety of settings.

Third grade to ninth grade correlations. The Direct Instruction groups were the only groups for which we could correlate a variety of end-of-third-grade scores to ninth grade reading and math scores. No comparable data were available for the control groups because individual student test scores are archived in New York City after ten years. It is most interesting to note that there are very high correlations (r = .78 and r = .81) between the spring third grade Metropolitan Achievement Test (the MAT) scores and the ninth grade Total Reading scores. This correlation is even higher than the correlation suggested by Bloom (1976) when he predicted that third grade and ninth grade standardized test scores would correlate +.75. This means that the end-of-third-grade MAT Total Reading score is a very good predictor of ninth grade reading success. Students who do well on the MAT at the end of third grade are highly likely to do well in ninth grade.

In comparison, the Wide Range Achievement Test (WRAT) third grade scores in reading and math correlate highly in reading (r = .47, r = .71, and r = .73) with ninth grade scores. These high correlations show the WRAT, too, to predict from moderately to very well from the end of third grade to ninth grade reading. The WRAT math test correlates moderately (r = .49 and r = .39) to ninth grade math scores, showing a positive relationship, though not an extremely high correlation.

Correlations of IQ and ninth grade scores. The Direct Instruction model administered the Slosson IQ test as a measure of language ability. The correlations are moderate for the Slosson on 9th grade scores for Cohort 1 and 2 (r = .37 reading, and r = .49 math; r = .20 reading and r = .49 math). This suggests simply that for whatever reason, IQ correlated more highly for Cohort 2 students in reading and math than it did for Cohort 1 students. The implementation of the Direct Instruction Model at P.S. 137 was equally strong for the Cohort 1 and 2 students, so there is no readily explainable difference in the correlations for these cohorts.

Implications and a suggestion for future research conclude this paper.

Implications

When one begins an intervention, one hopes for lasting change, particularly when intervening in an area that has the potential to effect subjects' lives. We must believe that some of the impetus for funding programs for the poor, such as Follow Through, as part of the War on Poverty stemmed from the desire to make life better on a short-term and long-term basis for disadvantaged minority people in the United States. In the late 1960's, federal support went to poor families and schools in the hope that increased funding and innovative programs would result in increased educational achievement and lifetime opportunities. This study has demonstrated that there are apparent long-term

effects of the Direct Instruction Follow Through program at P.S. 137. There are five major implications from this work.

First, these results are important in light of the criticism often leveled at the Distar Reading I and other synthetic phonics beginning reading programs. Critics of these approaches suggest that a reading program with an early emphasis on decoding will produce "word callers"-- students who decode words skillfully without comprehending what they read. Critics also contend that if students are first taught synthetic phonics, they may actually be hindered in the later grades when there is greater emphasis on comprehension. The data from this study suggest the contrary. The third grade and ninth grade tests are primarily tests of reading comprehension, not decoding. It appears that the Direct Instruction students are indeed comprehending what they read. The children who did well at the end of third grade are the same children who do well in ninth grade.

Second, there is recent evidence (New York Times, May 15, 1983) that minority students are dropping out of high school at rates higher than they were even three years ago. This problem is particularly grave in the inner cities. In New York City, for example, 45% of the students who enter ninth grade will drop out. A number of high school programs designed to interest students in staying in school have failed, and the focus is now returning, "to programs to raise student achievement long before the students get to high school" (Halverson, 1983). The results of

this longitudinal study strongly suggest that the roots of long-term success may well lie in an academic early childhood program. The die is cast by the end of third grade.

Third, the controversial Gates Program in New York (see Minter, 1982 for a description), or similar programs such as Mastery Learning in Chicago that emphasize student achievement of basic skills in the early grades and retaining students who fail to achieve mastery may well be on the right track. These programs operate on the philosophy that children should be close to grade level in the early grades and should not be allowed to move through school without achieving mastery. The findings from this study and the Lazar et al. (1977) work support early retention in mastery-based programs and suggest that perhaps it is particularly important to have students performing well by the end of third grade.

Fourth, regular elementary school programs can interpret these results to suggest that strong kindergarten through third grade programs with an academic focus on reading, language, and math with mastery may bolster disadvantaged students' chances for greater long-term academic success. In these days of decreased funds for education, studies such as this that strongly suggest that early interventions can have lasting effects should be considered seriously by policy-makers with smaller budgets and larger numbers of poor students. These results point to increased benefits for poor children who receive help in basic

skills when they enter school. Early intervention may be the most cost effective educational program on a long-term basis.

Fifth, the results of this study suggest that longitudinal research in education is important and worth the trouble. Only after collecting and analyzing these kinds of data can administrators, school board members, educators, researchers, and parents develop a sense of the importance and the effects of an academic early childhood education.

In their book, Early Schooling: Cognitive and Affective Outcomes, Entwisle and Hayduk (1982) suggest, "First, the early grades may be precisely the time that schools have their strongest effects...elementary schools tend to mirror the neighborhoods they serve. For this reason, school and home influences are probably most congruent at this life stage," (p. 136). Few would argue that in this country elementary schools tend to mirror their neighborhoods, and this study demonstrates that Entwisle and Hayduk may be correct that the early grades may be the time that schools have their strongest effects. The P.S. 137 students and the control students come from very similar neighborhoods, but they now have very different life chances. These groups received very different early childhood programs, and there have been significant long-term differences in the program's effects.

References

Becker, W. C., & Gersten, R. M. (1982). A follow-up of follow through: The later effects of the direct instruction model on children in fifth and sixth grades. American Educational Research Journal, 19, 75-92.

Bloom, B. (1976). Human characteristics and school learning. New York: McGraw-Hill.

Cook, T., & Campbell, D. T. (1979). Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally.

Darlington, R. B. (1981). The consortim for longitudinal studies. Educational Evaluation and Policy Analysis, 3(6), 37-45.

Durkin, D. (1978-79). What classroom observations reveal about reading comprehension instruction. Reading Research Quarterly, 14(4), 481-533.

Durkin, D. (1981). Reading comprehension instruction in five basal reader series. Reading Research Quarterly, 16(4), 515-544.

Engelmann, S. E., & Stearns, S. (1972). Distar Reading III. Chicago: Science Research Associates.

Entwisle, D. R., & Hayduk, L. A. (1982). Early schooling: Cognitive and affective outcomes. Baltimore: Johns Hopkins University Press.

Halverson, T. (1983). The New York Times, March 15.

Hunt, J. McV. (1982). Toward equalizing the developmental opportunities of infants and preschool children. The 1982 Kurt Lewin Address, University of Illinois at Urbana-Champaign.

Lazar, I., Hubbel, V. R., Murray, H., Rosche, M., & Royce, J. (1977). The persistence of preschool effects. Department of Health, Education, and Welfare, Publication No. (OHDS) 78-30130.

Maeroff, G. I. (1982). Don't blame the kids. New York: McGraw-Hill.

Mason, J. M. (1983). A question about reading instruction. In Duffy, Roehler, & Mason (Eds.), Comprehension instruction: Perspectives and suggestions. New York: Longman.

Meyer, L. A., Gersten, R. M., & Gutkin, J. (in press). Direct instruction: A project follow through success story. Elementary School Journal.

Minter, T. K. (1982). The importance of the federal role in improving educational practice: Lessons from a big city school system. Harvard Educational Review, 52(4), 500-513.

New York Times, June 20, 1982.

New York Times, March 15, 1983.

New York Times, "Dropout Rate 68% Hispanic Unit Says," June 10, 1983.

Schweinhart, L. J., & Weikart, D. P. (1980). Young children grow up: The effects of the perry preschool program on youths through age 15. Ypsilanti, MI: High Scope Educational Research Foundation.

Stebbins, L., St. Pierre, R. G., Proper, E. C., Anderson, R. B., & Cerva, T. R. (1977). Education as experimentation: A planned variation model, Vol. IVA-D. Cambridge, MA: Abt Associates.

Footnote

[1]Project Follow through was conceived by the Federal government in the mid 1960's as part of the War on Poverty. Twenty sponsors, early childhood university groups or educational labs, tested their programs throughout the United States. All approaches were comprehensive in that they had parent, health service, and educational components. The Direct Instruction Model implemented at P.S. 137 had a strong academic focus on basic skills for kindergarten through third grade.

Table 1

Demographic Comparison of the
Two Schools Prior to Follow Through

	Year	Reading[a] 2nd Grade Fall / Spring	Reading[a] 3rd Grade Fall / Spring	Ethnicity Black '69	Black '70	Black '71	Puerto Rican '69	Puerto Rican '70	Puerto Rican '71	% Low Income[b] (AFDC 1973)
P.S. 137	1966-1967	1.6 / 2.3	2.4 / 3.2	84.7	87.3	89.3	14.3	10.3	9.2	76.9
Control School	1966-1967	1.6 / 2.5	2.4 / 3.3	82.2	80.7	87.3	17.1	18.4	16.3	77.5

[a] Grade equivalents only are available for the Metropolitan Achievement Test (MAT) for these years. Fall/Spring scores were used as pretest/posttest measures for second and third grades.

[b] AFDC data were first kept by school in New York City in 1973. In both schools 100% of the students were eligible for free lunch, AFDC percentages were used to demonstrate a more rigorous test of socioeconomic need.

Table 2

Percentages of Cohort 1, 2, and 3 Students:
Graduated High School; Retained, Dropped Out;
Applied to College; and Accepted for College

	Cohort	Graduated	Retained	Dropped Out	Applied to College	Accepted for College
Direct Instruction	1	62%	4%	39%	50%	41%
Controls	1	38%	42%	62%	18%	9%
Direct Instruction	2	64%	50%	20%	23%	23%
Controls	2	38%	35%	58%	25%	21%
Direct Instruction	3	62%	10%	24%	38%	38%
Controls	3	39%	21%	18%	24%	21%

Table 3

T-tests for Aggregated Cohorts 1, 2, and 3 Students: Graduated, Retained, Dropped Out, Applied to College and Accepted for College

Category	Direct Instruction (N = 65)	Controls (N = 100)	$t_{(163)}$	p	Variance Accounted For
Graduated	59.5% (SD = 2.78)	37.6% (SD = 1.52)	58.12	<.001	97.4%
Retained	21.4% (SD = 24.9)	32.6% (SD = 10.69)	- 3.42	<.001	12.8%
Dropped Out	27.7% (SD = 10.07)	46% (SD = 24.3)	- 6.61	<.001	22.9%
Applied to College	34% (SD = 9.59)	18.5% (SD = 7.23)	11.14	<.001	53.0%
Accepted to College	34% (SD = 9.64)	17% (SD = 6.93)	12.3	<.001	58.8%

Table 4

T-tests for Cohorts 1, 2, and 3 and Aggregated

9th Grade Mean Reading and Math Scores in Grade Equivalents

Subject	Cohort	Direct Instruction/N	Controls/N	t	p	Variance Accounted For
Reading	1	9.6 (SD = 2.22) (N = 16)	7.59 (SD = 2.54) (N = 26)	$t_{(40)} = 2.609$	<.01	14.5%
	2	8.51 (SD = 2.16) (N = 17)	7.96 (SD = 1.92) (N = 30)	$t_{(45)} = .902$	<.37	1.8%
	3	9.4 (SD = 2.24) (N = 29)	8.94 (SD = 2.59) (N = 37)	$5_{(64)} = .759$	<.45	1.9%
	Aggregated 1, 2, 3	9.20 (SD = 2.22) (N = 60)	8.21 (SD = 2.42) (N = 93)	$t_{(151)} = 2.551$	<.01	4.1%
Math	1	8.3 (SD = 2.13) (N = 17)	7.79 (SD = 2.67) (N = 27)	$t_{(42)} = .667$	<.51	1.0%
	2	8.08 (SD = 2.52) (N = 17)	7.60 (SD = 1.79) (N = 29)	$t_{(44)} = .754$	<.46	1.3%
	3	9.03 (SD = 2.35) (N = 29)	8.33 (SD = 2.23) (N = 38)	$t_{(65)} = 1.244$	<.22	2.3%
	Aggregated 1, 2, 3	8.59 (SD = 2.33) (N = 61)	7.95 (SD = 2.23) (N = 93)	$t_{(152)} = 1.711$	<.09	1.9%

Table 5

Pearson Product Moment Correlations End-of-Third Grade
WRAT, Slosson, and MAT Scores to 9th Grade Reading and Math Scores

	Direct Instruction Students				
Cohort	9th Grade Rdg and 3rd Grade MAT Rdg	9th Grade Rdg and 3rd Grade WRAT Rdg	9th Grade Math and 3rd Grade WRAT Math	9th Grade Rdg and Slosson IQ	9th Grade Math and Slosson IQ
1	NA	.47	.49	.37	.20
2	.78	.71	.39	.49	.49
3	.81	.73	.39	NA	NA